COMPUTATIONAL VISION MODELS AND OCCUPATIONAL VISION STANDARDS

Bettina L. Beard and Albert J. Ahumada, Jr.
NASA Ames Research Center, Moffet Field, CA

Background: We describe a methodology that may be used to write uniform and universally accepted occupational vision standards. A simple image discrimination model is first calibrated using stimuli representative of airframe and powerplant cracks. It is then used to predict the visibility of simulated cracks of different lengths and widths. Visual acuity declines are simulated using a gaussian blur function on the crack images. Crack width is shown to be a salient cue to crack detection. Using this modeling technique we show when acuity declines begin to significantly effect performance. Future research will validate model predictions with human psychophysical data.

INTRODUCTION

In a recent review of the occupational vision standards literature, Beard et al. (2002) found that the majority of occupational vision standards are not empirically substantiated, and appear to be arbitrarily decided. A few standards have been empirically defined. For example, to define a visual acuity standard for police officers, Sheedy (1980) measured the size and working distance of the critical visual details for a representative task. Visual acuity standards have also been defined for police officers (Good, 1987; 1996), basket weavers (Good et al., 1996) and firefighters (Padget, 1989) using blurring lenses to reduce acuity while measuring performance on a job relevant task. Finally, Mertens et al. (2000) measured performance in color weak individuals on simulated ATC tasks to set an empirically defined color vision standard.

Currently no general standard exists in the aviation industry for the visual qualifications of maintenance inspectors. Some aircraft maintenance facilities have developed their own vision qualification programs, highlighting the need for a uniform and universally accepted set of vision standards that would apply to all aircraft non-destructive inspection and testing (NDI/NDT) personnel. It is difficult, if not impossible, to eliminate human error in the process of inspection. Therefore interventions must be developed to reduce these errors and make the process more error-tolerant. Since visual inspection represents 80% of all aviation maintenance inspection tasks (Goranson & Rogers, 1983), one mitigation strategy is to define vision standards for this vision-intensive, safety-critical occupation.

In this paper we apply a novel methodology toward defining an empirically based visual acuity standard for a representative task performed by aircraft maintenance personnel who do NDI/NDT and visual inspection. Computational models of human vision can make an important contribution to occupational vision requirements. One application of these models has been as image quality metrics, an application in which there are two images, an original image and a reconstructed version following image compression. The model predicts discriminability of the two images and thus the visibility of the compression artifacts (Watson, 1983). These discriminability models have also been used to predict object detection in a complex background, such as camouflaged military tanks (Rohaly et al.,

1997) and simulated aircraft on a runway (Ahumada & Beard, 1997).

To obtain an estimate of a visual acuity standard using image discrimination models, we follow a multi-step process. First, we calibrate the model for stimuli representative of airframe and powerplant cracks that are clear and blurred. We use a subset of the standard Modelfest images, whose contrast thresholds have been measure in a number of laboratories to calibrate the model. Second, we use the calibrated model to predict the visibility of simulated cracks of different lengths and widths as a function of blur, simulating reduced visual acuity in the image, rather than with blurring lenses, so that the image characteristics are exactly known. This provides an estimate of how much contrast sensitivity is lost by blur, so that if the tolerable loss in contrast sensitivity can be specified, the corresponding visual acuity is then specified. In support of the model's accuracy, we plan to obtain human psychophysical measurements to validate the simulated crack predictions. In addition, we will use the model to compare the simulated crack predictions to predictions for actual crack images in a natural aircraft scene. And finally, we will validate the natural scene predictions with human in the loop data. In this paper we report the results for the first two steps of this process.

The purpose of this paper is threefold. (1) To introduce a new methodology for determining occupational vision requirements. (2) To present the technique used for model calibration. (3) To run the model on simulated crack images over a range of widths and lengths at different levels of visual acuity.

METHODS & RESULTS

A Representative Defect

Aircraft inspection is a complex process, requiring many tasks, skills, and procedures. Its main purpose is the detection of discontinuities such as cracks[1] within the airframe and powerplant regions of the aircraft. Because these cracks may be very small and of low contrast, good visual acuity is likely to be involved in their detection. Visual acuity refers to a measure of spatial resolution of a person's vision for a high contrast, static image. After consulting with domain experts, we chose crack detection as the representative task in which to model.

A Simple Model

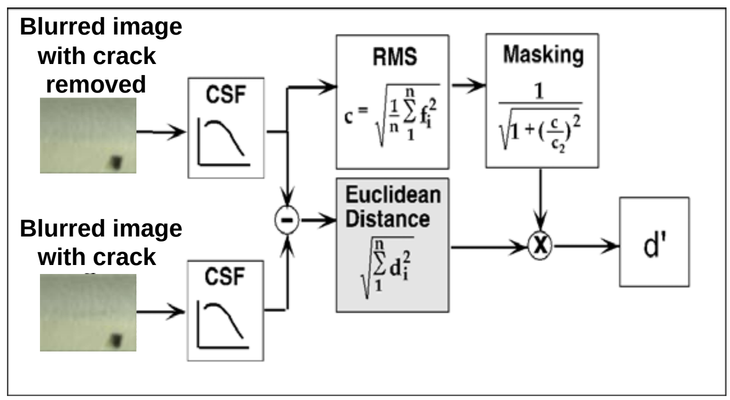

Figure 1. Schematic of an image detection model

Figure one's upper image is the background image and the lower image is the background-plus-defect image. The two input images (contrast images) enter the visual system, where they are filtered by a difference of gaussian blurring function. The difference of the images is calculated after which two standard deviations are computed; the first represents the root mean square error of the background image, which

[1] A crack may be defined as "A planar breach in continuity in a material" (Hellier, 2001). They are typically caused by two surfaces being overlaid at a boundary.

is assumed to be the masker and the second is the standard deviation of the defect pixel contrast. This generates a masking curve in which the masking contrast is determined by c^2. The product of these outputs represents the predicted sensitivity or the just noticeable difference of the crack defect.

Image discrimination models predict the difference in visibility between two similar images. The models take two images as input, and output a prediction of the number of Just Noticeable Differences (JNDs) between them. In this version of the model, one luminance image is considered to be a blurred version of the background image and the other is the blurred background-with-crack image. These images are filtered using the Contrast Sensitivity Function (CSF) in order to normalize sensitivity. The model takes the contrast energy in the target and adjusts it by the background variance.

Model Calibration

To provide a common data set for the development of models of contrast target detection, the Modelfest project developed a set of 44 images, most of which are various grating patches (the entire set of 44 calibration images can be obtained from *http://vision.arc.nasa.gov/modelfest*). To calibrate our model, we chose seven of the 44 images because of their physical similarity to aircraft crack defects. These seven images are shown in Figure 2.

Earlier predictions of real world stimuli (Rohaly et al., 1997; Ahumada & Beard, 1997) have assumed a contrast sensitivity function (CSF) with a sinusoidal grating threshold of 1%. To fit the average (n=16) Modelfest thresholds for the stimuli in Figure 2 we need to use a best grating threshold of 0.5%. We tried Minkowski summation exponents of 2 and 4 and found that the best fit for these seven stimuli was a summation exponent of 2 (Euclidean Distance). When the entire set of 44 images was run through the model, the best fitting exponent was 4 (probability summation). This is probably because many of the other images in the set of 44 contained extended, high spatial frequency features whereas the seven images used here either were localized within a small spatial area or contained only extended low frequency energy.

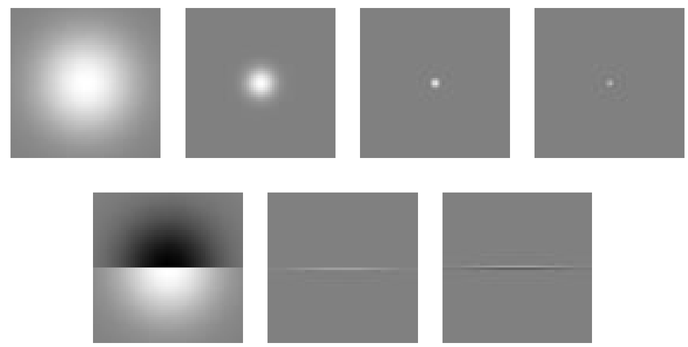

Figure 2. Stimuli used to calibrate the contrast discrimination model. The leftmost 4 images are Gaussians with decreasing standard deviations, the fifth through seventh images are an edge, line, and dipole respectively.

Simulating Visual Acuity Decline

Although the shape of the human blur function differs between individuals and changes for different optical conditions, it can be approximated by a Gaussian spread function. The model has a difference of Gaussians contrast sensitivity function with a center Gaussian spread of 2 min. To simulate different levels of visual acuity, we blur the image with a Gaussian and then report the acuity as the ratio of the effective center spread to the original model value. Thus we are assuming that the model has 20/20 vision. For example, if the blur has a spread of 2 min, the effective center Gaussian spread will be root 2 times 2 min (Pythagorean rule) so that the effective acuity will be 20/28.

Model Predictions

We next predicted the visibility of a set of simulated cracks as a function of blur (simulating visual acuity declines) for a range of lengths and widths. The widths were 0.5, 1, 2, 4, and 8 min. The lengths were the widths times 1, 2, 4, 8, and 16. Figure 3 shows how the threshold contrast for each image varied as a function of blur relative to the threshold for the unblurred image. The top curve is the result for the pinpoint crack (e.g., 0.5 min x 0.5 min). The threshold for this image is more affected by blur than the threshold for any other image. The figure shows that if the allowed sensitivity degradation were 6 dB (a factor of 2 in contrast), the allowable acuity degradation would be about 20/60.

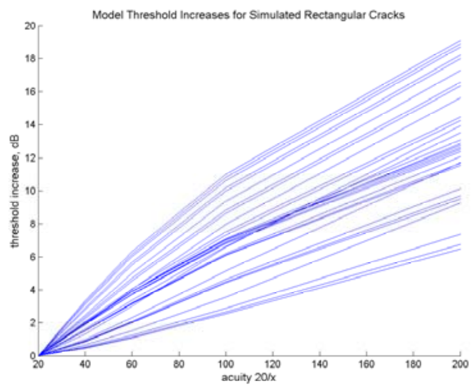

Figure 3. Increments in contrast thresholds in dB as a function of visual acuity decline for the range of crack length and widths described in the text. The top curve is for the smallest crack (0.5 min by 0.5 min), the bottom curve is for the biggest crack (8 min by 128 min).

DISCUSSION

The first aim of this paper was to describe a methodology that may be used to generate empirically based occupational vision standards. It does not provide a standard, but it converts the problem to specifying a desired physical limitation in performance. Here we use this technique to help define the spatial vision requirements for aircraft NDI/NDT personnel using simulated crack images. These modeling results will help define the parameters tested in the human psychophysical experiments. We next need to validate that line detection predicts actual aircraft crack detection.

Vision is a fundamental component of effective aircraft maintenance inspection. All the same, so too are other cognitive factors such as attention, memory, and experience. Inspectors are knowledgeable about individual components as well as the overall aircraft being inspected, thus they possess the background to properly locate, identify, and evaluate aircraft defects. Therefore, although vision is a critical component in inspection, other factors weigh in heavily on the naturalistic task.

ACKNOWLEDGMENTS

Supported by FAA Agreement DTFA-2045. Thanks go to Dr. Willa Hisle for her extensive review of the occupational standards literature and for her comments on this paper.

REFERENCES

Ahumada Jr., A.J., Beard, B.L. (1997). Image discrimination models predict detection in fixed but not random noise. *Journal of the Optical Society of America A* **14**, pp. 2471- 2476.

Beard, B.L., Hisle, W.A., Xing, J., & Ahumada, A.J. (2002) Behind the setting of vision standards for aircraft maintenance inspection. Report for FAA AAR-100.

Good, G. W. & Augsburger, A. R. (1987). Uncorrected visual acuity standards for police applicants. *Journal of Police Science and Administration*, 15(1), 18-23.

Good, G. W., Weaver, J. L., &

Augsburger, A. R. (1996). Determination and application of vision standards in industry. *American Journal of Industrial Medicine*, 30(5), 633-630.

Hellier, C. J. (2001) *Handbook of Nondestructive Evaluation*. New York: McGraw-Hill.

Mertens, H. W. & Milburn, N. J. (1996). Performance of color-dependent air traffic control tasks as a function of color vision deficiency. *Aviation, Space, and Environmental Medicine,* 76, 919-927.

Padgett, V. R. (1989). Summary of empirical validation of firefighter vision requirements: four field experiments. Paper presented at the International Personnel Management Association Assessment Council 13th Annual Conference held in Orlando, FL. Retrieved July 11, 2002, from http://home.earthlink.net/~medtox/firevis.html

Sheedy, J. E. (1980). Police vision standards. *Journal of Police Science and Administration*, 8(3), 275-285.

Watson, A.B. (1983) Detection and recognition of simple spatial forms. In O. J. Braddick and A.C. Sleigh, eds., *Physical and Biological Processing of Images*, pp. 100-114, Berlin: Springer-Verlag.